Journal of Speech Sciences 3(2): 1-20.2013
Available at: www.journalofspeechsciences.org

FOREIGN ACCENT: LENGTH OF RESIDENCE, LISTENER AND STIMULUS EFFECT

WANG, Xinchun.
California State University

Abstract

Mandarin speakers' productions of English sentences, spontaneous speech, and filtered speech were rated for degree of foreign accent by native English and Mandarin listeners. Results showed Mandarin speakers with 12 years' length of residence (LOR) in the U.S. were rated to be accented as those with zero LOR. Untrained native Mandarin listeners with no LOR in the target language environment were comparable to native English listeners in gauging degree of foreign accent based on sentences and spontaneous speech. No stimulus effect was found between sentences and spontaneous speech for accent rating. Filtered natural speech appeared to attenuate degree of foreign accent and Mandarin listeners were not able to assess foreign accent based on long excerpts of filtered speech. The findings suggest that LOR is not an important predictor of degree of foreign accent for adult speakers with late age of arrival (AOA).

Key words: foreign accent; length of residence; accent rating; listener effect, stimulus effect

---

*Corresponding author: xinw@csufresno.edu

1. Introduction

It is well-known that adult second language (L2) learners often speak the target language with a noticeable foreign accent, which is defined as "non-pathological speech that differs in some noticeable respects from native speaker pronunciation norms" (Munro & Derwing, 1995). Although foreign accented speech can be highly intelligible, it may take a longer time to process and more effort to comprehend (Constance & Merrill, 2004; Floccia, Butler, Goslin, & Ellis, 2009). Researchers have often pointed out that foreign accent, intelligibility (how well the speech is actually understood), and comprehensibility (how difficult it is to understand the speech) are related but independent factors that are measured differently in L2 speech studies (Munro & Derwing, 1999). Despite these differences, heavy foreign accent often affects intelligibility and comprehensibility.

Previous research has examined different factors that influence adult L2 speakers' degree of foreign accent. The single most important factor has been consistently found to be the learners' age of arrival (AOA) in the target language country, the beginning point of first-time, sudden, and massive exposure to the target language (Evers, Reetz, & Lahiri, 1998; Flege, 1988; Flege et al., 2006; Flege & Fletcher, 1992; Flege, Munro, & MacKay, 1995; MacKay, Flege, & Imai, 2006; Munro & Mann, 2005; Piske, MacKay, & Flege, 2001; Trofimovich & Baker, 2006). Other factors have also been explored for their role in L2 speakers' degree of foreign accent. These factors include (but are not limited to) the speaker's first language (L1), age of learning (AOL), length of residence in the L2 environment (LOR), the amount of L1 and L2 use, formal instruction and phonetic training in L2, aptitude, and motivation (Derwing & Munro, 2005; Derwing & Rossiter, 2003; Moyer, 1999; Munro & Derwing, 2001; Piske et al., 2001). It is important to point out that in the Internet and international travel age, physical AOA and LOR as variables may not indicate the beginning point of heavy exposure to the target language as they used to indicate in previous studies. Yet, controlled studies on communication with native speakers via the Internet and its impact on foreign accent have not been reported in the literature. Future studies need to examine this new phenomenon.

This study examines the effect of LOR on degree of foreign accent on two groups of highly advanced Mandarin learners of English, who teach in the target language and who share many important learner characteristics but differ in LOR (zero vs. 12 years). The main goal is to explore whether 12 years of LOR in North America provides any advantage for highly advanced L2 English speakers in terms of reduced degree of foreign accent. A secondary goal is to examine the listener effect on foreign accent rating. A group of native Mandarin speakers who have never been abroad but who learned English as a foreign language in China will rate native Mandarin and native English speakers' English productions for degree of foreign accent. A third goal is to examine the differences, if any, between different stimulus types (read sentences, spontaneous speech, and filtered speech) on foreign accent rating. The review of the

literature in the following sections focuses on LOR, the listener effect and stimulus effect on foreign accent rating.

1.1 Length of Residence

The effect of LOR on degree of foreign accent has been found to be inconsistent. One of the reasons for the mixed findings might be the differences in the LOR gap between the L2 groups being examined (Piske et al., 2001). For example, Trofimovich & Baker (2006) found that native Korean speakers with a mean of 10 years of LOR in the U.S. were rated by native English listeners to be as accented as those with three years of LOR. However, both groups were rated less accented than the inexperienced group with only three months of LOR. The results suggest that an extremely short LOR of three months in the native speaking environment is not sufficient for L2 speakers to reduce foreign accent. Interestingly, additional 7 years of LOR (LOR 3 years vs. 10 years) did not appear to give the native Korean speakers any advantages in terms of reduced degree of foreign accent. Similarly, Flege (1988) found that foreign accent ratings obtained for two groups of native Mandarin speakers with different LOR (1.1 vs. 5.5 years) in the U.S. were not significantly different. Based on some of these findings, researchers hypothesized that after a rapid initial phase of learning, LOR may not affect the pronunciation of individuals who began learning L2 as adults (Piske et al., 2001).

The effect of LOR on degree of foreign accent has also been found to be related to the learning stages and the speakers' L2 proficiency levels. Research has shown that additional years of LOR may not help to change the degree of L2 foreign accent for highly experienced learners (Piske et al., 2001) while it does appear to make a difference for speakers at early stages of L2 learning (Riney & Flege, 1998; Riney, Takagi, & Inutsuka, 2005).

It is important to note that previous studies on the effect of LOR on degree of foreign accent often compared an experienced group with longer LOR with an inexperienced group with shorter LOR in the target language countries. Learners' English as a Foreign Language (EFL) experience with the target language in non-native speaking environment is usually not well controlled for foreign accent research. One exception is the study by Bongaerts and colleagues (Bongaerts, van Summeren, Planken, & Schils, 1997) on highly successful native Dutch speakers of English whose productions of English sentences were rated as native-like by native English listeners. These Dutch speakers had only one year of LOR in Britain at some point in the process of L2 learning but the authors speculated that their native-like accent might be attributed, at least in part, to the speakers' intensive training (unspecified) in L2 pronunciation. Their daily use of L2 in teaching the target language in Dutch universities might also be a factor. More studies are needed to investigate these factors and their impact on the degree of foreign accent, ideally, with other L1 speakers. Experienced L2 speakers with zero LOR but who use the target language daily in EFL situations

have not been included in such studies. The current study will fill this gap by examining the effect of LOR on the degree of foreign accent with two groups of native Mandarin learners of English who differed in LOR (zero vs. 12 years) but were comparable in other aspects of L2 experience.

1.2 Foreign Accent Judgment

In L2 speech studies, the overall degree of foreign accent is often assessed by listeners through rating tasks using equal-interval rating scales. Research has shown that both the speech being judged and the language experiences of the listeners play a part in assessing degree of foreign accent. It is obvious that certain nonnative intonation and stress patterns or certain mispronounced segments are easily identified by the native listeners as foreign accented (Magen, 1998). Temporal features such as speech rate and pause also affect the degree of foreign accent. These factors are "speaker-dependent" (Levi, Winters, & Pisoni, 2007).

On the other hand, listeners' first language experience, in particular, their experience with the accent of the speakers' first language may also affect the judgment of degree of foreign accent. This "listener effect" is independent of the speech being rated because the listeners may bring their own language experience to the task of gauging the degree of foreign accent. For example, if the listeners share the same first language with the speakers being judged, or, if the listeners are familiar with the speakers' native language, they may be more tolerant with their ratings of the speakers' L2 foreign accent than those listeners who are not familiar with the speakers' L1. In an L2 intelligibility study, such shared L1 listener benefit was reported as "matched interlanguage speech intelligibility benefit" (Bent & Bradlow, 2003).

In contrast, Munro, Derwing, and Morton (2006) found that there was a significant degree of shared experience for the four listener groups (Cantonese, Japanese, Mandarin, and English) when they judged the speech produced by L2 speakers from linguistic backgrounds similar to and different from their own. In particular, the listeners did not consistently exhibit an intelligibility benefit for speech produced by speakers who shared their own L1. Similarly, MacKay et al. (2006) also found that native English listeners and native Arabic learners of English did not differ in their judgment of degree of foreign accent on native Italian speakers' productions of English sentences. The high degree of correlation of both native and nonnative listeners' ratings indicate that nonnative speakers are just as capable as native English speakers in gauging degree of foreign accent.

Flege (1988) found that native Tawainese speakers with different length of residence (5.1 years vs. 1.1 year) in the U.S. did not differ in degree of foreign accent in their production of English sentences as judged by native English listeners. However, the more experienced group with longer LOR was more similar to the native English listener group in judging the nativeness of Taiwanese speakers' productions of English sentences. The findings suggest that nonnative listeners' ability to gauge foreign accent was

influenced by their LOR in the target language environment even though their degree of foreign accent was not judged to be different by native English listeners.

It is important to point out that previous studies on listener effect normally involved ESL learners who themselves had been in the target language environment with years of LOR. Listeners who are EFL learners and who have never lived in the target language environment are not commonly included in foreign accent judgment experiments. This study will fill this gap by investigating whether nonnative listeners who are EFL learners with zero LOR are capable of gauging degree of foreign accent.

1.3. Stimulus Effect and Foreign Accent Rating

Another factor that may impact the listeners' judgment of foreign accent is the stimulus effect. Read sentences are commonly used for foreign accent rating assessment (Flege, 1988; MacKay et al., 2006; Munro et al., 2006; Riney and Flege, 1998; Riney et al., 2005) and very few previous studies included different stimuli for comparisons. The level of extemporaneity (read speech vs. spontaneous speech) and the length of the stimulus (e.g. sentence vs. paragraph) may both influence foreign accent rating. For example, Munro and Mann (2005) tested the effect of length of the stimuli on perceived degree of foreign accent by comparing words, sentences, and paragraphs elicited through reading tasks (read speech with different length). They also tested the effect of extemporaneity by comparing a read paragraph and free narrative speech elicited through description of pictures. The results suggested that the ratings of nativeness decreased with sampling length (read speech) but increased with extemporaneity (comparing read and extemporaneous speech.) They concluded that sentences offer the best tool for judging foreign accent when taking into consideration the trade-off between the length and extemporaneity of the stimulus being rated.

In contrast, Bongaerts, Planken, & Schils (1995) found that sentences received the least native-like and words the most native-like ratings among the four different stimuli: Recount of a recent trip abroad, sentences, a short paragraph, and words (Bongaerts, Planken, & Schils, 1995). There were no differences between the narratives and the read texts. Similarly, Moyer (1999) found no significant differences between the use of sentences, read paragraphs, and narratives, but isolated words received more native-like ratings than all the other stimulus types (Moyer, 1999). These different findings on stimulus effect in previous studies suggest that the influence of stimulus on foreign accent rating was not consistent.

Modified speech used for accent rating is also common in foreign accent studies. For example, read speech at normal rate can be edited to increase or decrease the speech rate to test whether fast or slow speech affects nonnative speakers' degree of foreign accent (Munro & Derwing, 1995, 2001). Other modified speech, such as low-pass filtered speech, has been used to reduce the impact of segmental errors

by removing the segmental information but retaining the suprasegmental features of the speech. For example, Munro (1995) found that native English listeners were able to rate the degree of foreign accent based on low-pass filtered unintelligible English sentences produced by L2 speakers. Because these low-pass filtered sentences retained only their temporal and suprasegmental features, the listeners were provided with the contents of the sentence in writing to ensure that they knew what sentence they were rating (Munro, 1995).

The current study investigates the use of low-pass filtered natural utterances (longer than the sentence) to reduce the effect of possible non-phonological errors on foreign accent rating. It also examines the differences, if any, in the use of read sentences and natural speech in foreign accent judgment.

1.4. The Current Study

This study examines the effect of LOR on degree of foreign accent on two groups of highly advanced Mandarin learners of English who teach in the target language and who share many important learner characteristics (see Table 1) such as years of L2 learning and age of first exposure to L2 but differ in LOR (zero vs. 12 years). The main goal is to explore whether 12 years of LOR in North America provides any advantage for highly advanced L2 English speakers in terms of reduced degree of foreign accent. Both native Mandarin speaking groups consist of university professors who teach in the target language and use English as a second language (in the U.S) or as a foreign language (in China) on a daily basis. Their productions of sentences and spontaneous speech as well as low-pass filtered speech will be rated for degree of foreign accent.

A secondary goal is to examine the listener effect in foreign accent rating. A group of native Mandarin speakers who have never been abroad but learned English as a foreign language in China will rate native Mandarin and native English speakers' English productions for degree of accent. Their rating results will be compared with those of native English listeners. Previous research that examined nonnative listeners' ability in gauging foreign accent often involved ESL learners who had been immersed in the target language environment. The listeners in the current study are naïve EFL learners who had no target language experience in an ESL environment. A third goal is to examine the differences, if any, between the effects of different stimulus types on foreign accent rating. The research questions are:

1) Will native Mandarin speakers with 12 years of residence in North America be rated less accented in English than those controlled for other L2 experience but with zero LOR?

2) Will native Mandarin listeners who have never lived abroad be able to judge native Mandarin speakers' L2 English foreign accent in the same way the native English listeners do?

3) Are listeners' judgments of degree of foreign accent influenced by different stimulus types including read sentences, spontaneous speech, and low-pass filtered speech?

2. Method

2.1. Speakers

Two groups of highly advanced native Mandarin learners of English participated as speakers. They were the American Professor group (the AP group) and the Chinese Professor group (the CP group). The AP group consisted of 10 professors (6 male, 4 female, mean age = 43.4, range 33-52) teaching at a university in the United States at the time of the study. They were all born and raised in China and studied English as a foreign language for a mean of 10 years (5-20) at school and earned at least one university degree in China before they moved to North America. Their mean age of learning (AOL), age at which learning English as a foreign language began at school, was 14.3. One of the AP speakers started learning English at the age of five (at home) and another at 20 (at college), and the remainder began learning English between the ages of 10-12. Their mean age of arrival (AOA) in North America was 31.3 (23-37). Their mean length of residence in North America was 12.4 years (5- 21). All had earned at least one university degree in North America. They had taught at universities in the United States for a mean of. 12.4 years (2-18). Their subject areas of teaching are science, engineering, business, education, and humanities. As expected, the AP group reported a very high mean percentage use of English, 97.5% (90%-100%) outside home but a mean of only 37% (5%-95%) use of English at home.

The CP group consisted of 10 professors (4 male, 6 female, mean age = 40.3, range 33-59) from two universities in China. Their mean age of learning (English as foreign language at school in China) was 13.1. One of them began learning English at college at the age of 23. All the rest began learning English at school between the ages of 10 to 14. All had at least one university degree in English or English linguistics from China but none had earned a degree from an English speaking country. They had taught English or English linguistics in universities in China for a mean of 13.2 years (4-25). Two speakers reported having taken short business trips overseas but none had lived abroad by the time of data collection for this study in 2005. None of the CP participants reported using the Internet to chat with native English speakers. Their mean use of English at home was 2.5% (0%-10%). Only three participants reported speaking English at home (5% - 10%) and they reported using it mainly when helping their children with English homework. None of the speakers reported speaking English at home regularly for any other purpose. However, they spoke English exclusively in class while teaching and used English to prepare their English and linguistics lessons. The background information of the AP and CP groups is presented in Table 1. A series of one-way ANOVAs revealed no significant differences between the AP and CP groups on any variables except on L2 use at home and outside home. Five native English speakers

who were professors in a U.S. university, the EP group (3 male, 2 female, mean age =40) participated as a control group. They were all from the same university from which the AP group was recruited.

Table 1. The AP and CP group speakers' background information

	Gender	Age	AOL	Year* learn	Year* teach	Self* Rate	Use E Outside	Use E Home	AOA	LOR
AP01	F	36	12	9	10	3	95	95	23	13
AP02	M	52	20	7	15	3	100	50	36	15
AP03	M	51	10	10	18	5	100	10	30	21
AP04	M	41	11	11	12	5	100	20	26	15
AP05	M	50	10	10	15	5	100	10	35	19
AP06	F	47	12	12	22	5	100	20	37	10
AP07	M	43	12	7	8	7	100	20	35	8
AP09	F	40	12	10	13	3	90	90	35	5.2
AP10	M	41	10	5	9	5	100	5	31	10
AP11	M	33	5	20	2	4	90	50	25	8
	Mean	43.4	11.4	10.1	12.4	4.5	97.5	37.0	31.3	12.4
CP01	M	42	14	11	10	5	In Class	0		
CP02	M	36	12	12	6	5	In Class	10		
CP03	M	59	13	11	25	4	In Class	0		
CP04	F	34	10	13	9.5	6	In Class	5		
CP05	F	49	12	9.5	22	5	In Class	0		
CP06	M	34	10	13	11.5	4	In Class	0		
CP07	F	33	10	14	11	4	In Class	10		
CP08	F	29	13	13	4	5	In Class	0		
CP09	F	53	23	3	27	3	In Class	0		
CP10	F	34	14	12	6	4	In Class	1		
	Mean	40.3	13.1	11.15	13.2	4.5		2.6		

Year learn: number of years speakers have studied English.
Year teach: number of years speakers have taught.
Self rate: speakers' self rating of accent along a scale of 1(native-like) – 9 (heavy accent)

2.2. Stimuli

The stimuli used for rating tasks were sentences, spontaneous speech, and low-pass filtered spontaneous speech. The sentences were elicited through a reading task in which the speakers read a list of 10 short sentences at normal speed. Only two of the ten, one statement and one question, were used for rating: "Most people like to listen to music." "Do you have a gas cooker in your kitchen?" Spontaneous speech was elicited through an interview in which the speaker answered two related questions: "Do you enjoy teaching?" and "What do you like the most about teaching?" The speakers heard the questions from the researcher twice and were given 1- 2 minutes to prepare before they provided the answers. To avoid

the questions-and-answers format, and to elicit extemporaneous speech that is natural and uncontrolled, the speakers were asked to express their opinions about teaching and relate their experience to the topic. The recordings were made in a quiet room using a Sanyo micro cassette recorder (TRC-680MN) and an external microphone. The recorded sentences and interviews were digitized at a sampling rate of 22050 Hz with 16-bit resolution and normalized for peak intensity. For the spontaneous speech (hereafter referred to as utterances), a short excerpt of a single continuous sample was extracted from the beginning of each speaker's speech using waveform editing. The length of the utterances ranged from 10 -16 seconds. To produce a unit that sounded like a completed turn, all utterances were cut at the end of a syntactic boundary, accompanied by falling intonation. Each selected excerpt of spontaneous speech was also low-pass filtered to preserve the suprasegmentals such as stress, rhythm, and intonation by removing the segmental information. The cut-off frequency was 350Hz for female speakers and 250Hz for male speakers using MLS software. All test stimuli were recorded onto an audio CD for presentation.

2.3. Listeners

The native English listeners were 11 monolingual English speakers (4 male, 7 female, mean age = 27.6, range 21-37) recruited from the student population at the same university where the AP and EP groups were recruited. According to self-report, all native English listeners were familiar with foreign accents and mostly with Spanish accents. However, none of them were familiar with Mandarin accented speech. All reported having normal hearing. A small honorarium was paid to the English listeners for the rating tasks.

The Mandarin listeners were unpaid volunteers. They were 18 undergraduate students (mean age = 19.6, range 18-22) recruited from a university in China. They had learned EFL since elementary school and were taking English courses at the university at the time of this study. According to self-report, none of the native Mandarin listeners had been abroad. The listeners in neither group were familiar with the speakers whose speech they rated.

2.4. Procedure

The total stimuli used for accent rating were 108 sentences (2 sentences × 27 speakers × 2 presentations), 27 utterances, and 27 filtered utterances. The listeners rated the stimuli in the order of sentences, utterances, and filtered speech in three separate blocks. In the first block, the two sentences were presented in two separate sub-blocks. Each sentence was randomized and presented with one repetition. In the utterance block, the 27 utterances were also randomized but each utterance was presented once without repetition. In the filtered speech block, the same 27 utterances were low-pass filtered, randomized, and presented once.

Before the rating task began, the researcher explained each listening block to the listeners to make sure that they fully understood the tasks. This was followed by a practice session in which the listeners practiced the rating tasks with sample sentences, sample utterances, and sample filtered speech produced by both native and nonnative speakers. These samples were not included in the data analysis. The native English listeners performed the rating tasks in small groups of 3-5 in a sound-treated language lab where the audio CD was played through a built-in speaker system at comfortable level. The listeners rated each item they heard on an answer sheet by circling a number on a Likert scale of 1 (native-like) to 9 (heavy foreign accent). They were encouraged to use the full scale. The native Mandarin listeners performed the same rating tasks in the same manner the native English listeners did, except that the stimuli were played through a loud speaker in a quiet classroom.

Twenty-five out of 27 speakers' data were analyzed. One English speaker's data were excluded for analysis due to problems with the clarity of her speech. Another AP speaker's data were also eliminated because it was later found that her LOR was too short to meet the minimum requirement of five years.

3. Results

3.1. Reliability

Interrater reliability and intraclass correlations for the two listener groups' ratings on the 25 speakers' sentences, utterances, and filtered speech were calculated. Interrater reliability (α range .74 - .98) and intraclass correlations (.70 - .96) ranged from moderate to very high (see Table 2), which suggests the members within the English listener group (N=11) and the Mandarin listener group (N=18) tended to agree with one another in their judgments. Therefore, each speaker's rating scores for sentences (the average of the two sentences), utterances, and filtered speech were calculated by taking the mean of all the listeners in each listener group.

Table 2. Interclass and intraclass reliability of ratings of Sentence, Utterance, and Filtered Speech by native English listeners (N = 11) and native Mandarin listeners (N = 18)

English listeners						Mandarin listeners					
Utterance		Sentence		F. Speech		Utterance		Sentence		F. Speech	
inter	intra	inter	intra	inter	intra	inter	intra	inter	intra	inter	intra
.975	.964	.961	.946	.735	.703	.868	.860	.893	.881	.743	.732

3.2. Results of Speaker and Listener Groups

The mean accent ratings for the AP, CP, and EP speaker groups' sentence, utterance, and filtered speech given by the English and Mandarin listener groups are presented in Figure 1 and Figure 2 respectively. A two-way repeated measures ANOVA was carried out on the mean ratings of the speakers by both listener groups. The between-subject factor was speaker group (AP, CP, EP) with listener group (English, Mandarin) and stimuli (Sentence, Utterance, Filtered Speech) as within subject factors. This analysis revealed a significant main effect of speaker groups $F(2,22) = 42.69$, $p = .000$. Post hoc pairwise comparisons revealed the differences between the EP and AP groups, and between the EP and CP groups were significant ($\alpha < .01$). The differences between the AP and CP groups were not significant.

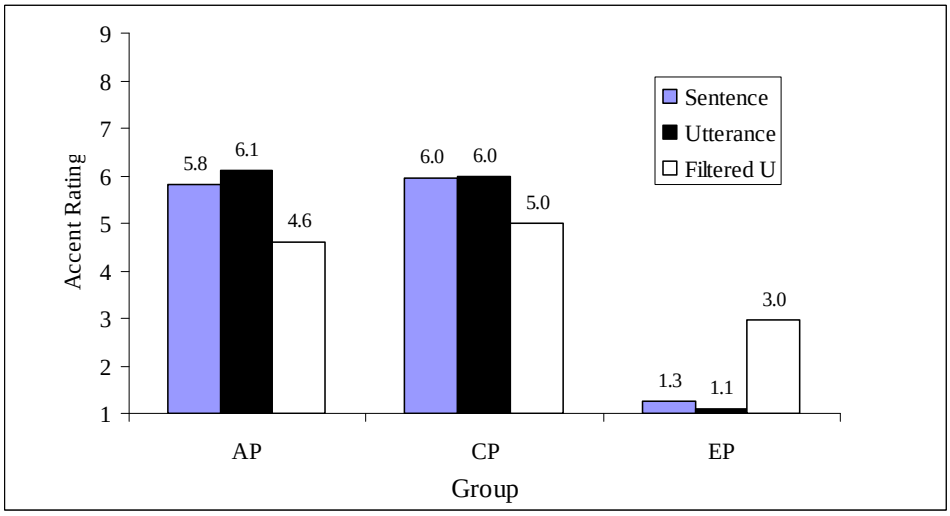

Figure 1. The AP, CP, and EP groups' mean accent ratings given by English listeners

The factors of listener $F(1,22) = .39$, $p = .54$ and stimuli $F(2,22) = .22$, $p = .81$ were not significant. Nor was the listener × stimuli interaction $F(2,22) = 2.8$, $p = .07$. However, the listener × speaker interaction $F(2,22) = 32.59$, $p = .000$, the stimuli × speaker interaction, $F(4,44) = 9.07$, $p = .000$, as well as the listener × stimuli × speaker interaction $F(4,44) = 3.86$, $p = .01$ were significant. These interactions suggest that the listener groups differed in their ratings of the three speaking groups on some stimulus types. To further investigate these differences, a series of one-way between group ANOVAs were carried out on native English and Mandarin listener groups' ratings separately.

One-way between group ANOVAs for the English listener group's mean ratings revealed significant differences between the speaker groups on sentence $F(2,24) = 55.08$, $p = .000$, on utterance $F(2,24, = 74.95$, $p = .000$, and on filtered speech $F(2,24) = 8.67$, $p = .002$. Post hoc Bonferroni test (adjusted for multiple comparisons, $\alpha < .01$) established significant differences between the EP and AP

groups and between the EP and CP groups on all three stimulus types. The differences between the AP and CP groups were not significant on any of the three stimulus types.

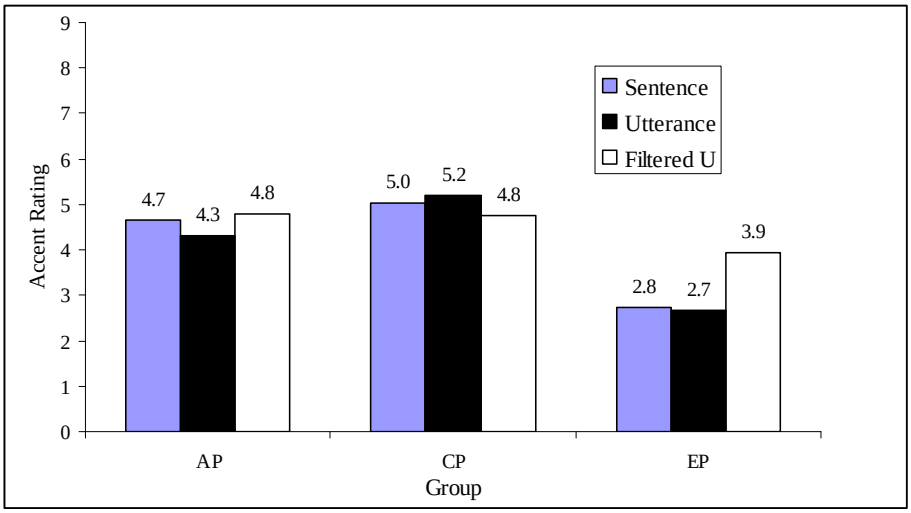

Figure 2. The AP, CP, and EP groups' mean accent ratings given by Mandarin listeners

The same one-way between group ANOVAs for Mandarin listener group's ratings established significant differences between the speaker groups on sentence $F(2,24) = 9.30$, $p = .001$ and on utterance $F(2,24, = 13.01$, $p = .000$. Post hoc Bonferroni test established significant differences ($\alpha < .01$) between the EP and AP groups and between the EP and CP groups but not between the AP and CP groups. The one-way ANOVA on filtered speech by Mandarin listener group was not significant $F(2,24) = 1.907$, $p = .172$. These analyses suggest that the interaction between listener, speaker, and stimuli was due to the differences between the two listener groups in their judgment of the filtered speech across the three speaker groups. The Mandarin listeners failed to tell the native and nonnative differences based on the filtered speech.

3.3. Individual Speakers

Each individual speaker's mean accent rating scores given by the native English listeners are presented in Table 3. An inspection of these data revealed that most Mandarin speakers' mean rating scores of sentences and utterances fall between 5 and 7 along the scale of 1-9. The two speakers (AP11 and CP07) who had the lowest scores (most native-like) in the AP and the CP groups respectively received mean rating scores of 4-5. These scores were far above the native English speakers' mean ratings of 1.1 and 1.3 for sentences and utterances respectively. In fact, not a single AP or CP group member obtained a mean rating score that fell within 2 standard deviations of the native English speakers' mean score, a standard often used to judge whether a nonnative speaker's speech is native-like or not (Bongaerts et al., 1997; Flege, 1988).

As shown in Table 3, individual speakers in the AP and CP groups are comparable in their degree of foreign accent regardless of the dramatic differences in the LOR. However, the mean ratings on filtered speech for the AP and CP groups were not consistently different from the EP group. Two AP group speakers (AP02 and AP10) received a mean rating score of 3.1, which was comparable to native EP group mean of 3.

Table 3. Individual speakers' mean rating scores of Sentence, Utterance, and Filtered Speech given by native English listeners (N = 11)

AP	Sent.	Utter.	F.SP	CP	Sent.	Utter.	F.SP	EP	Sent.	Utter.	F.SP
AP01	6	5.3	5.1	CP01	6.8	6.4	4	EP01	1.1	1	2.9
AP02	6.1	6.6	3.1	CP02	7.4	6.7	3.9	EP02	1.2	1.3	3.4
AP03	6.4	6.5	5	CP03	7.2	6.5	5.8	EP03	1.3	1.1	2.4
AP04	5.5	5.5	4.5	CP04	6.5	7.1	5.3	EP04	1.5	1.1	2.5
AP05	5.7	6.1	4.6	CP05	5.3	5.7	4.5	EP05	1.1	1	3.7
AP06	5.1	6.9	5.7	CP06	4.1	3.8	4.5				
AP07	7	7.4	3.9	CP07	5.1	5.4	4.5				
AP09	5.2	5.7	6.2	CP08	5.8	5.4	4.9				
AP10	6.7	6	3.1	CP09	6.8	7.3	6.5				
AP11	4.4	5.2	4.7	CP10	4.7	5.5	6.1				
Mean	**5.8**	**6.1**	**4.6**		**6.0**	**6.0**	**5.0**		**1.2**	**1.1**	**3.0**

4. Discussion

This study investigated the effect of LOR on degree of foreign accent as well as the listener and stimulus effect on foreign accent rating. The results suggest that 12 years of LOR did not appear to give the AP group any advantages in terms of reduced degree of foreign accent compared with the CP group with zero LOR. Mandarin listeners not residing in the target language environment were comparable to native English listeners in gauging degree of foreign accent. With regard to stimulus effect, no difference was found between read sentences and spontaneous speech for accent rating. Low-pass filtered spontaneous speech appeared to attenuate degree of foreign accent for native English listeners. Mandarin listeners were not able to gauge degree of foreign accent based on filtered spontaneous speech.

4.1. Length of Residence and Foreign Accent

The current data did not provide support for previous findings of positive effect of LOR on perceived degree of foreign accent. For example, Flege et al. (1995) examined various factors such as age of learning, gender, L1 and L2 use, and LOR that contributed to native Italian speaker's degree of foreign accent in their productions of English sentences as judged by native English listeners. They found the

Italian speakers' 15 -32 years of LOR in Canada to be small but significant variance in foreign accent rating. In another study, Trofimovich & Baker (2006) found that experienced native Korean ESL speakers with longer LOR (3-10 years) in North America were rated less accented than the inexperienced speakers with only three months of LOR. Their results suggest that extremely short LOR of three months in the target language environment was not sufficient for foreign accent reduction while longer LOR did make a difference. It is important to note that the current study examined highly advanced L2 speakers who used the target language daily in academic environment while the native Korean speakers did not share the same L2 experience. The EP group in the current study arrived in America much later than the Italian speakers (ranged between ages of 2-23). Therefore, the findings may not be directly comparable to those of Flege et al. (1995) and Trofimovich & Baker (2006) studies. These differences lead to the question of whether differences in L2 speakers' proficiency level upon their arrival in native speaking countries influence their L2 phonological improvement and degree of foreign accent.

The current data also show that none of the AP and CP group speakers was rated as native-like. These results are in agreement with previous research, which often led to the conclusion that adult L2 speakers were rarely rated as native-like in the target language, regardless of their extended LOR (Flege et al., 1995). One exception is the Bongaerts et al. (1997) study on highly successful native Dutch speakers of English whose productions of English sentences were rated as native-like by native English listeners. Even though the AP group in the current study had much longer LOR (12 years) in North American than the native Dutch speakers in Britain (1 year), the Dutch speakers had many advantages that the AP group did not have. Dutch is much more closely related to English in terms of its phonological system than Mandarin Chinese is, in particular, its prosody. Furthermore, the native Dutch speakers had one year of residence in Britain much earlier in life than the AP group members who were in their late 20s and 30s when they arrived in North America. Additionally, it is likely that the Dutch speakers had more access to the target language environment through trips to Britain than the AP group members who did not have this opportunity earlier in life before they moved to North America.

Factors that influence the perceived degree of foreign accent are complicated as the current data suggest. Age of learning, age of arrival, the amount of L2 use, learning stages, and L2 proficiency all have been found to play a part in the speakers' perceived degree of foreign accent (Piske et al., 2001). Social factors such as attitudes and motivation may also influence L2 learners' degree of foreign accent.

For the current study, the focus was on LOR and the related factors such as L2 use and AOA. In terms of L2 use, although both the CP and AP groups spoke the target language exclusively in an academic setting as professors on a daily basis, the CP group did not speak English in non-instructional activities in their other daily activities. None of them travelled to English speaking countries regularly and the majority of them had not been abroad prior to this study. None reported using the Internet to chat with

native English speakers. In contrast, the AP group used English more than the CP group because they resided in the U.S. and had to speak English outside the classroom in their daily life in addition to their teaching activities. Despite these differences in L2 use and LOR, the AP group did not outperform the CP group in reduced degree of foreign accent.

Previous research has consistently indicated that age of arrival is the most important factor that influences degree of foreign accent (Flege et al., 1995; Munro & Mann, 2005; Piske et al., 2001). Young children with early AOA in native speaking countries were often rated as native-like while adult speakers who arrived later in life were not judged as native like although they have comparable length of residence (Flege et al., 2006; Flege et al., 1995; Munro & Mann, 2005). The AP group arrived in the U.S. as adults with a mean age of 31 years, long after they had become very fluent L2 speakers. It appears that LOR is not an important predictor for reduced degree of foreign accent for advanced L2 learners with late AOA.

L2 speakers' first exposure to the type of input at early stages of learning may be another important factor that contributes to a learner's L2 phonological development and foreign accent. The participants in both AP and CP groups grew up in China before it opened its door to the outside world and international travel was rare for its citizens. Both the AP and CP groups were exposed to the target language at an early age through classroom instruction in the EFL environment. Such limited exposure to the target language input mainly from their nonnative English teachers' instructions in the foreign language classroom might have influenced the learners' L2 phonology and foreign accent very early on in life. It is likely that an adult speaker's overall global foreign accent, once established early on in life in the EFL environment, becomes difficult to change even after a significant length of residence in the target language environment.

4.2. Listener Effect

Another goal of this study was to investigate the listener effect on perceived degree of foreign accent. Mandarin listeners not residing in the target language environment were able to rate degree of foreign accent as the native English listeners did on sentences and spontaneous speech. These findings are in agreement with previous research (Munro et al., 2006). One important contribution of this study was that the native Mandarin speaking listeners in the current study were EFL learners with zero LOR while most previous studies involved ESL listeners residing in the target language countries. The current findings suggest that LOR is not a required condition for foreign accent rating, especially when the listeners share the same L1 with the speakers. It appears that L2 learners' ability in rating degree of foreign accent far exceed their ability to produce accent-free target language. The relationship between L perception and production is not straightforward. Previous studies on L2 perception and production mainly focused on L2 segments (Best, 1995). It would be interesting to see future studies investigate the

differences between L2 perception and production on degree of foreign accent.

4.3. Stimulus Effect

An additional goal of the current study was to examine the stimulus effect on foreign accent judgment. The results showed that read sentences and spontaneous speech did not make a difference in foreign accent rating and these findings are in agreement with some previous studies (Moyer, 1999; Munro & Mann, 2005). The current data suggest that differences in degree of spontaneity between read sentences and free speech are not important for foreign accent judgment.

The results on ratings of filtered speech were mixed, as the native Mandarin listeners were not able to tell the differences between native and nonnative productions based on the unintelligible filtered speech. One reason for their inability to rate filtered speech for degree of foreign accent might be the inherent difficult task of rating long stretch of muffled speech with no lexical information. A close examination of the native English listeners' rating data on filtered speech may shed some light on the level of difficulty of this task. Although their overall rating scores of the filtered speech were not statistically different from their ratings of the sentences and utterances, the native English listeners were not as successful in rating filtered speech. On a scale of 1 to 9, their ratings of differences between the native and nonnative productions on sentences and spontaneous speech were around 5, while the difference was under 2 for the filtered speech. This difference also put their ratings of some of the AP and CP group members' filtered speech within the mean rating of the EP group members. In contrast, the English listeners did not rate any Mandarin speakers' productions of sentences and utterances as native-like.

It is important to point out that the filtered speech for accent rating in the current study was spontaneous speech with a mean length of 10-16 seconds. This is very different from previous research in which only filtered sentences were rated and the listeners were provided with the written content of the actual sentence they were rating (Munro, 1995). The listeners would be able to predict the prosodic features such as the intonation pattern and the rhythm of that particular sentence while rating because they knew the content. In contrast, the rating of lengthy filtered spontaneous speech is much more demanding even for the native English listeners as they had no knowledge of the content of the speech they were rating. The findings suggest that filtered speech longer than a sentence should be used with caution for accent rating.

5. Conclusions

In conclusion, extended length of residence in the U.S. did not give the AP group any advantage in terms of reduced foreign accent. The findings suggest LOR is not an important predictor for reduced degree of foreign accent for advanced L2 learners with late AOA. The current data have provided further

evidence in the literature for highly advanced learners who used the target language in academic settings. The highest attainment of native-like, accent free speech for adult learners is not affected by the LOR as much as by other factors.

Nonnative speaking listeners who learned English as a foreign language with zero LOR were able to gauge degree of foreign accent. Though less rigid than the native English listeners in judging Mandarin speakers' L2 English productions of sentences and spontaneous speech, the native Mandarin speaking listeners' ratings were comparable to those of native English listeners. Therefore, residence in the target language environment is not a required condition for nonnative speakers to accurately judge degree of foreign accent.

Read sentences and spontaneous speech did not appear to contribute to differences in foreign accent rating. Low-pass filtered speech may be used with caution when the stimulus is beyond the length of a sentence.

REFERENCES

Bent, T., & Bradlow, A. R. The interlanguage speech intelligibility benefit. The Journal of the Acoustical Society of America. 2003;114:1600-1610.

Best, C. T. A direct realist view of cross-language speech perception. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research. Baltimore: York Press; 1995. p.171-204.

Bongaerts, T., Planken, B., & Schils, E. Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z. Lengyel (Eds.), The Age Factor in Second Language Acquisition. Clevedon, UK: Multilingual Matters; 1995. p.30-50.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. Age and ultimate attainment in the pronunciation of a foreign language. Studies in Second Language Acquisition. 1997;19:447-465.

Constance, M. C., & Merrill, F. G. Rapid adaptation to foreign-accented English. The Journal of the Acoustical Society of America. 2004;116:3647-3658.

Derwing, T. M., & Munro, M. J. Second language accent and pronunciation teaching: A research-based approach. TESOL Quarterly. 2005;39:379-397.

Derwing, T. M., & Rossiter, M. J. The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. Applied Language Learning. 2003;13:1-17.

Evers, V., Reetz, H., & Lahiri, A. Crosslinguistic acoustic categorization of sibilants independent of phonological status. Journal of Phonetics. 1998;26:345-370.

Flege, J. E. Factors affecting degree of perceived foreign accent in English sentences. Journal of the Acoustical Society of America. 1988;84:70-79.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. Degree of foreign accent in

English sentences produced by Korean children and adults. Journal of Phonetics. 2006; 34:153-175.

Flege, J. E., & Fletcher, K. L. Talker and listener effects on degree of perceived foreign accent. Journal of the Acoustical Society of America.1992; 91:370-389.

Flege, J. E., & Liu, S. The effect of experience on adults' acquisition of a second language. Studies in Second Language Acquisition. 2001; 23:527-552.

Flege, J. E., Munro, M. J., & MacKay, I. Factors affecting strength of perceived foreigh accent in a second language. Journal of the Acoustical Society of America. 1995; 97:3125-3134.

Floccia, C., Butler, J., Goslin, J., & Ellis, L. Regional and foreign accent processing in English: Can listeners adapt? Journal of Psycholinguistic Research. 2009;38:379-412.

Levi, S. V., Winters, S. J., & Pisoni, D. B. Speaker-independent factors affecting the perception of foreign accent in a second languagea. Journal of the Acoustical Society of America. 2007;121:2327-2338.

MacKay, I. R. A., Flege, J. E., & Imai, S. Evaluating the effects of chronological age and sentence duration on degree of perceived foreigh accent. Applied Psycholinguistics. 2006;27:157-183.

Magen, H. S. The perception of foreign-accented speech. Journal of Phonetics. 1998;26:381-400.

Moyer, A. Ultimate attaninment in L2 phonologg: The critical factors of age, motivation, and instruction.

Studies in Second Language Acquisition. 1999;21:81-108.

Munro, M., & Mann, V. Age of immersion as a predictor of foreign accent. Applied Psycholinguistics. 2005;26:311-341.

Munro, M. J. Nonsegmental factors in foreign accent: Ratings of filtered speech. Studies in Second Language Acquisition. 1995;17:17-34.

Munro, M. J., & Derwing, T. M. Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. Language and Speech. 1995;38:289-306.

Munro, M. J., & Derwing, T. M. Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. Studies in Second Language Acquisition. 2001; 23:451-468.

Murno, M. J., Derwing, T. M., & Morton, S. L. The Mutual Intelligibility of L2 Speech. Studies in Second Language Acquisition. 2006;28:111-131.

Piske, T., MacKay, I. R. A., & Flege, J. E. Factors affecting degree of foreign accent in an L2: A review. Journal of Phonetics. 2001;29:191-215.

Riney, T. J., & Flege, J., E. Changes over time in global foreign accent and liquid identifiability and accuracy. Studies in Second Language Acquisition. 1998;20:213-243.

Riney, T. J., Takagi, N., & Inutsuka, K. Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. TESOL Quarterly. 2005;39:441-466.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. Studies in Second Language Acquisition. 2006;28:1-30.